Rendering Volumetric Haptic Shapes in Mid-Air using Ultrasound

Benjamin Long Sue Ann Seah Tom Carter Sriram Subramanian

Department of Computer Science, University of Bristol, UK*

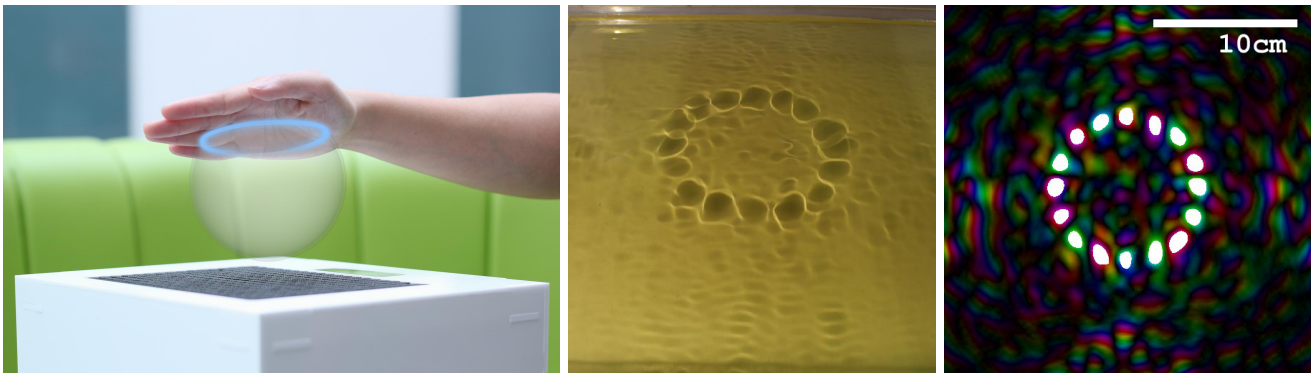

Figure 1: *Left: Ultrasound is focused to create the shape of a virtual sphere. Middle: Ultrasound pushes the impression of a sphere section into oil. Right: A simulation of the ultrasound at the plane intersected by the hand, where a threshold has been applied to highlight the foci.*

Abstract

We present a method for creating three-dimensional haptic shapes in mid-air using focused ultrasound. This approach applies the principles of *acoustic radiation force*, whereby the non-linear effects of sound produce forces on the skin which are strong enough to generate tactile sensations. This mid-air haptic feedback eliminates the need for any attachment of actuators or contact with physical devices. The user perceives a discernible haptic shape when the corresponding acoustic interference pattern is generated above a precisely controlled two-dimensional phased array of ultrasound transducers. In this paper, we outline our algorithm for controlling the volumetric distribution of the acoustic radiation force field in the form of a three-dimensional shape. We demonstrate how we create this acoustic radiation force field and how we interact with it. We then describe our implementation of the system and provide evidence from both visual and technical evaluations of its ability to render different shapes. We conclude with a subjective user evaluation to examine users' performance for different shapes.

CR Categories: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Haptic I/O.

Keywords: 3D haptic shapes, tactile displays, acoustic radiation forces

Links: DL PDF

*e-mail:{b.long,s.a.seah,t.carter,sriram.subramanian}@bristol.ac.uk

The definitive version was published in ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia), Volume 33, Number 6, Article 181, Publication Date: November 2014.

http://doi.acm.org/10.1145/2661229.2661257

1 Introduction

Haptics is a key factor in enhancing presence if we are to create fully immersive and realistic virtual environments [Reiner 2004]. Haptic feedback has been developed and evaluated for many applications ranging from teleoperation, entertainment, rehabilitation and even surgical training [Hayward et al. 2004]. This diverse use of haptic feedback has led to numerous methods of producing it in virtual reality and augmented reality systems.

One of the most common techniques for exploring a virtual environment is using proxy devices such as the SensAble by Phantom or the Falcon by Novint. The SensAble PHANTOM allows the user to interact with a virtual scene by moving a pen through 3D space [Bianchi et al. 2006]. When the pen comes into contact with a virtual object, an articulated arm attached to the pen provides resistance, thereby enabling the user to feel it. This method is also possible with magnetic levitation, removing the need for the articulated arm [Berkelman et al. 1996]. However, these techniques are not ideal for see-through and reach-through displays where images are floating in space. Attachments could be worn on the users' hand e.g. data gloves [Wusheng et al. 2003; Dipietro et al. 2008], but this requires the process of fitting them on before use. This can be cumbersome and prevents instantaneous user interaction.

An elegant solution would be to produce tactile sensations in mid-air where users are able to get haptic feedback without contact with any physical object or any actuator attached. Research has shown that this can be achieved using air jets [Suzuki and Kobayashi 2005], air vortices [Sodhi et al. 2013], and ultrasound [Carter et al. 2013; Hoshi et al. 2010]. Among these three, we have identified ultrasound as the most flexible and dynamic method for producing volumetric haptic shapes.

In this paper, we present a method for rendering volumetric haptic shapes using focused ultrasound. These haptic shapes can be experienced by the user with their bare hands, giving them the ability to 'walk-up and use' as they do not need any tools or attachments as shown in Figure 1.

Our main contributions are:

1. We present the first non-contact haptic feedback system capable of producing feelable three-dimensional shapes.
2. We improve upon existing algorithms to render large numbers of control points in real time with a predictable frame rate.
3. We introduce three optimizations to increase the strength of the tactile region.
4. Through technical and user studies we demonstrate that our algorithm works and users can accurately identify 3D shapes without visual feedback.

2 Background and Related Work

2.1 Mid-Air Haptic Devices

There have been a number of methods to provide haptic feedback in mid-air but so far none of them have been used to produce shapes. AIREAL [Sodhi et al. 2013] is one such technology that uses directable air vortices to create tactile sensations in 3D space. Here, the tactile sensations are of low fidelity as the area of stimulation is large and there is a latency to producing the air vortices in a specific location. Air jets have also been used, but these lack accuracy and are difficult to control [Suzuki and Kobayashi 2005]. A method which has a higher fidelity is using ultrasound-based acoustic radiation forces [Hoshi et al. 2010; Alexander et al. 2011; Carter et al. 2013], which produces multiple individually perceivable points of 1 cm diameter in mid-air. We advance on this method to be able to create volumetric shapes.

2.2 Creating Tactile Sensations using Focused Ultrasound

The use of focused ultrasound as a non-invasive method to stimulate neuroreceptor structures in various parts of the human body has been a topic of research since the early 1970s [Gavrilov and Tsirulnikov 2012]. Dalecki et al. [1995] first proposed the idea of using water-based ultrasound to create tactile sensations on a finger attached to an acoustic reflector floating at the surface of a water bath. It was later demonstrated that the skin can itself act as an acoustic reflector in a medium of air, thus realising the use of air-based focused ultrasound to produce tactile sensations [Carter et al. 2013; Hoshi et al. 2010].

These tactile sensations are caused by a non-linear effect of focused ultrasound called *acoustic radiation force*. The radiation force induces a shear wave in the skin tissue, creating a displacement, which triggers the mechanoreceptors within the skin [Gavrilov and Tsirulnikov 2002]. The maximum displacement u_{max} of a medium induced by radiation force from a pulse of focused ultrasound is defined by Gavrilov and Tsirulnikov [2002] as:

$$u_{max} = \begin{cases} \frac{\alpha a}{\rho c_l c_t} t_0 I, & \text{where } t_0 \ll a/c_t \\ \frac{\alpha}{\rho c_l c_t^2} a^2 I = \frac{\alpha}{c_l \mu} a^2 I = kW, & \text{where } t_0 \gg a/c_t \end{cases} \quad (1)$$

where a is the radius of the focal region, t_0 is the duration of the pulse, c_t is the speed of the shear waves propagation, c_l is the speed of sound, μ is the shear elastic modulus, α is the absorption coefficient, I and W are the intensity and acoustical power (both averaged over the pulse duration) and k is an amalgamated constant.

The tactile sensations however cannot be felt continuously unless it also changes continuously with time, as tactile receptors are mainly sensitive to changes skin deformation roughly between 200Hz and 300Hz [Gescheider and Wright 2008]. Thus, the ultrasound has to be modulated at a frequency which corresponds to the peak sensitivity of the tactile receptors.

2.3 Multiple Focal Points using Two Dimensional Phased Arrays

Two dimensional phased arrays of ultrasound transducers enable tactile sensations to be produced in three dimensions in mid-air. Hoshi et al. [2010] describes a system using a linear focusing method to dynamically create and move a single focal point. They also suggested a Fourier transform based inverse technique, but this would be fundamentally limited to a single plane of feedback parallel to a well-sampled plane of transducers.

Based on their method of generating a single focal point, Alexander et al. [2011] created up to four focal points by spatial multiplexing (treating subsections of the array as separate arrays to create single focal points) and temporal multiplexing (reconfiguring the array to produce single focal points in different places serially in time). Both these methods suffer from either the secondary maxima of multiple focal points constructively interfering with each other thus creating extra regions of perceivable haptic feedback or conversely the residual ultrasound from a focal point destructively interfering with other focal points.

Carter et al. [2013] proposed a solution to this problem by introducing the concept of null control points, at which the amplitude of the ultrasound is minimized. Any secondary maxima can then be then eliminated by positioning a null control point on them. This solution is an adaptation of a focusing method proposed by Gavrilov [2008]. Both of these techniques are based on Ebbini and Cain [1989] creating multiple simultaneous control points, wherein a minimum norm step containing an explicit inversion (which is time-consuming and can generate numerical instability) has been augmented with a weighting matrix. This minimum norm step containing the weighting matrix is then iterated to convergence in order to achieve maximum power output for a given control point configuration.

Even though Hertzberg et al. [2010] optimised Gavrilov's algorithm and improved its efficiency with regards to maximising transducer power, his algorithm still took more than 70 ms to find a solution with only 9 control points. As this technique uses multiple iterations to a convergence criterion, the run time can fluctuate and a smooth frame rate will be difficult to achieve. This implies that the solutions previously described are not suitable for rendering volumetric shapes.

When interacting with virtual content, a hand will make contact with several parts of an object at once. It is therefore necessary to provide three dimensional haptic feedback all across the hand which is fast enough so that it is not perceived as discontinuous. To create a volumetric shape we will require a far larger number of control points and an even faster run time. We will also need to regulate the amplitude of each control point and ensure that the array is efficient at generating these control points. To achieve these ends, our work obviates the costly matrix inversion and the iterative reweighting method used in previous work, replacing it with a different formulation that is more robust, predictable and efficient on modern parallel hardware. We describe how these changes are effected in the following section.

3 Algorithm: Controlling an Acoustic Field

An array of ultrasonic transducers can be described as a collection of apertures emitting sound waves of a known frequency. Using

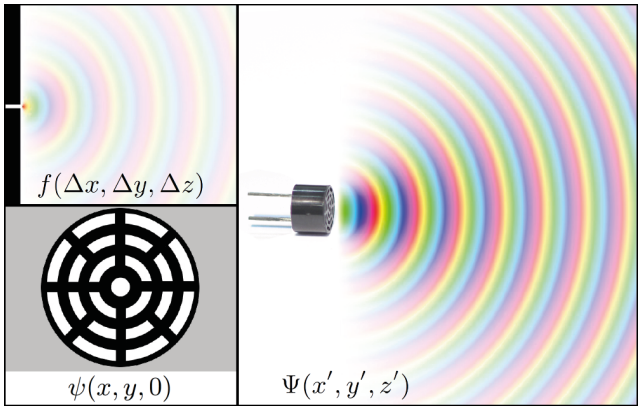

Figure 2: *An illustration showing the field functions that make up the Rayleigh-Sommerfeld diffraction integral. Top left: The diffraction convolution function can be interpreted as the result of the wave passing through an infinitesimal slit. Distance from the slit Δz increases to the right. Bottom left: The aperture (shown in white) that the planar wave passes through, in this case the grille from the upper surface of the transducer. Right: The convolved function, approximating the wave field emitted from the single transducer shown. Again, distance from the aperture z' increases to the right.*

the Rayleigh-Sommerfeld diffraction integral we can relate known sound wave phases and amplitudes traveling through the aperture with spatially defined phases and amplitudes in the far field [Gavrilov 2008].

In Section 3.1, we descibe how to synthesize an acoustic field as the solution to a linear system. By modelling the output of an ultrasonic transducer and approximating the near field behavior, fast algorithms can be used to compute the numerical integrations involved by expressing the diffracted sound wave function from a single transducer. This can then be used to characterize a single transducer so that we can solve for a transducer configuration that closely reproduces a set of field values at given points. As the produced acoustic field is controlled only at these points, they are known as 'control points'.

In Section 3.2 we describe an algorithm to calculate a control point phase that interferes with nearby control points in a way that induces amplitude gain. For the purposes of haptic feedback, controlling the amplitude at each sampled control point is important, but the phase is not. We can thus choose an appropriate phase for each point that interferes constructively with other points that have high amplitude requirements and destructively with other points that must be low in amplitude.

It is also possible that a desired configuration of control point amplitudes, specified by a phased-array focusing technique, can be generated with large variances in amplitudes at the source transducers. This is unwanted, as running transducers at low power results in weaker phenomenon, while using too much power can damage the array elements. As most ultrasonic arrays power all transducers at the same level, ignoring amplitude, this introduces artifacts when the amplitude recommendations made by a phased-array focusing technique are normalized away and not followed. We show in Section 3.3 that power demand variance can be penalized so that powering the array at full does not cause unwanted artifacting and detail deterioration.

In order to produce haptic feedback we must modulate the ultrasound at a perceptable frequency (as described in Section 2.2). Due

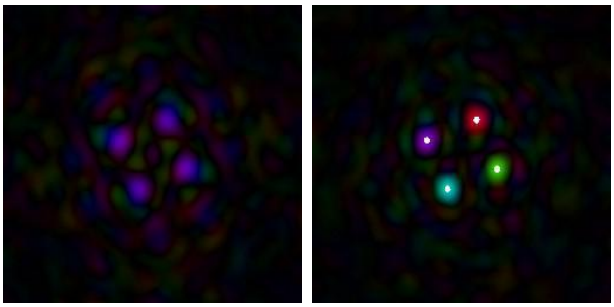

Figure 3: *When control points are created they have a local residual field. By modifying the phases of the control points, this field can be exploited to apply gain to all control points. This is the function of the eigensystem solver. On the left, we show a set of four control points which have defaulted to the same phase setting. On the right, the eigenvector encodes a set of phases that results in added gain to all control points.*

to this, the transducers have to be modulated by emitting ultrasound for half of the time, while being powered down for the remaining duration. This is inefficient, and so we describe in Section 3.4 a technique for splitting the output into multiplexed parts to effectively increase the strength of the array.

We then evaluate the algorithm by producing and comparing simulations and impressions on the surface of liquids to validate our implementation. Finally, we show the performance of the algorithm and its behavior with increasing numbers of control points.

3.1 Waveform Synthesis Algorithm

To build a model of the acoustic field Ψ generated by the n ultrasonic transducers, we assume each transducer emits a planar wave at an aperture conceptualized as a two-dimensional wavefunction ψ. The Rayleigh-Sommerfeld integral [Ebbini and Cain 1989] gives the far field behavior of a two-dimensional wavefunction diffracting through an aperture of known geometry. It can be expressed in three-dimensions as:

$$\Psi(x', y', z') = \frac{-\Delta z}{i\sqrt{\lambda}} \iint_{\Omega} \psi(x, y, 0) \cdot f(\Delta x, \Delta y, \Delta z)\, dx\, dy,$$
$$f(\Delta x, \Delta y, \Delta z) = \frac{e^{ik\sqrt{(\Delta x)^2+(\Delta y)^2+(\Delta z)^2}}}{\left((\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2\right)^{3/4}}, \quad (2)$$

in which

$$\Delta x = x - x', \qquad \Delta y = y - y', \qquad \Delta z = z - z' \quad (3)$$

where x, y and z define coordinates relative to the aperture, while x', y' and z' define absolute positions in the far field, as illustrated in Figure 2. In our case, we define $\psi(x, y, 0)$ to give a circular surface of constant phase and unit amplitude.

The functional form of f permits this to be expressed as a convolution for each slice in z', giving:

$$\Psi(x', y', z') = \psi(x, y, 0) \otimes f_{z'}(x, y) \quad (4)$$
$$= \iint \psi(x, y, 0) \cdot f_{z'}(-\Delta x, -\Delta y)\, dx\, dy \quad (5)$$

which can be accelerated using the fast Fourier transform (FFT) algorithm [Nascov and Logofătu 2009]. With this, we can generate a look up table that describes how the amplitude and relative phase of the sound wave from one transducer changes spatially.

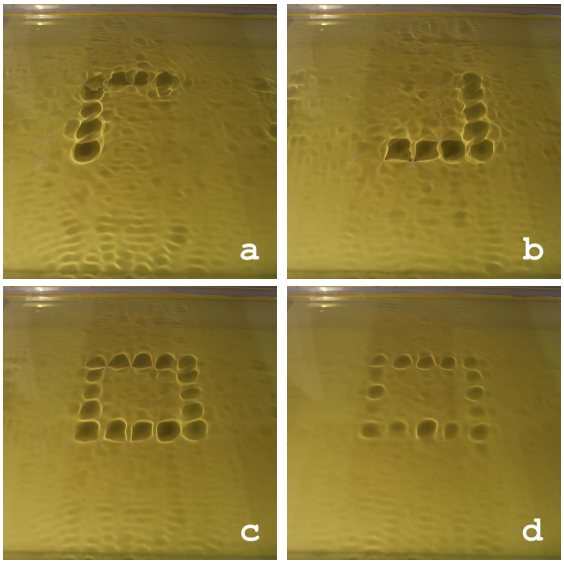

Figure 4: *By using both parts of the modulation period, complex modulated output can be effectively doubled in power by multiplexing two acoustic fields. a) The left hand side of the presented feedback. b) The right hand side of the presented feedback. c) The two fields switching at the modulation frequency required for haptic perception and generating a powerful result. d) The result when modulating between a single acoustic field and an unpowered state. The square shown is approximately 10 centimetres across.*

For a transducer q, the sound wave can now be split into four parts: the product of an emission amplitude A_q^{emit}, a phase offset e^{ϕ_q}, and in the far field, an amplitude attenuation function $A_q^{\text{attn}}(x', y', z')$ and a phase difference function $e^{ik_q(x',y',z')}$. The latter two of these are dealt with by the diffraction look-up function, as each of the n transducers are the same:

$$\begin{aligned}\Psi_\Omega(x', y', z') &= \sum_{q=1}^{n} A_q^{\text{emit}} e^{\phi_q} \cdot A_q^{\text{attn}}(x', y', z') e^{ik_q(x',y',z')}, \\ &= \sum_{q=1}^{n} A_q^{\text{emit}} e^{\phi_q} \cdot \Psi_q(x', y', z'), \end{aligned} \tag{6}$$

At this point we can choose a set of m control point positions in x', y' and z', which we will denote $\{\boldsymbol{\chi}_1, \ldots, \boldsymbol{\chi}_m\}$, attributing to each a complex number describing phase and amplitude and assert that these are part of the field $\Psi_\Omega(x', y', z')$ generated by the ensemble of ultrasonic transducers. The emission amplitudes and phase offsets required to produce these can then be obtained by specifying the necessary simultaneous equations as a complex-valued $\mathbf{Ax} = \mathbf{b}$ linear system, where:

$$\mathbf{A} = \begin{bmatrix} \Psi_1(\boldsymbol{\chi}_1) & \ldots & \Psi_n(\boldsymbol{\chi}_1) \\ \vdots & \ddots & \vdots \\ \Psi_1(\boldsymbol{\chi}_m) & \ldots & \Psi_n(\boldsymbol{\chi}_m) \end{bmatrix}, \tag{7}$$

the vector $\mathbf{x} = [\, A_1^{\text{emit}} e^{\phi_1}, \ldots, A_n^{\text{emit}} e^{\phi_n} \,]^T$ and $\mathbf{b} = [\, \Psi'_\Omega(\boldsymbol{\chi}_1), \ldots, \Psi'_\Omega(\boldsymbol{\chi}_m) \,]^T$. Once in this form, $A_1^{\text{emit}} e^{\phi_1}$ is a complex coefficient that can be solved for using either an under-determined minimum norm solver or over-determined least squares formulation. Such a formulation solves for the optimal set of initial amplitude and phase offsets that the transducers are required to emit in order to produce the desired control points.

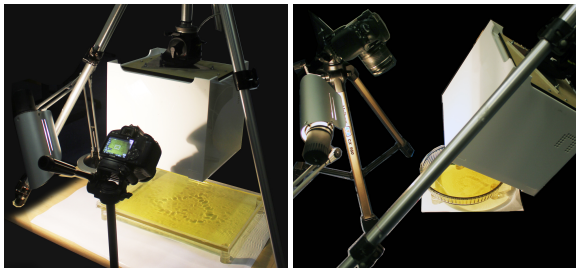

Figure 5: *Our setup for capturing the impressions of acoustic fields on the surface of oil. We also set up our system to render test shapes at a incline, as shown in this set up.*

However, we are not as interested in reproducing a set of desired amplitude and phase point measurements as we are in reproducing amplitudes for the purpose of producing a fast tactile sensation. To create a variety of sensations we need to create multiple areas with both high and low amplitudes to provide a contrast and to control noise. To achieve this, we must first determine a set of compatible phases that can co-exist above the array at the desired amplitudes.

3.2 The Control Point Position Phase Eigenproblem

In order to create the set of desired amplitudes at given positions, we must first determine what phases relative to each other the control points must have to most effectively take advantage of local constructive and destructive interference. Because the complex-valued acoustic field smoothly changes in space, the spatial and amplitude relationships between control points also has repercussions for phase relationships among control points which can be chosen to amplify or dampen the output. For example, control points at the same amplitude will amplify each other as they move closer if the phase difference is zero, or can alternatively dampen and cancel as the distance narrows if the phase difference is half a period. This is also true for the local residual field surrounding each control point, as is shown in Figure 3.

We consider a good candidate for the phases of the n control points to be represented by the x-vector in an eigenproblem $\mathbf{Rx} = \lambda\mathbf{x}$. We choose the matrix $\mathbf{R}$ to represent the phase shift and amplitude effects that an efficient solution for each individual control point has on each of the others. Given that we can quite simply find a solution for any one control point, we then use symbolic algebra to algebraically generate a simplified minimum-norm solution for each single control point case:

$$A_q^{\text{emit}} e^{i\phi_q} = \frac{A_q^{\text{attn}}(\boldsymbol{\chi}_\mathcal{C} - \boldsymbol{\chi}_q) e^{ik_q(\boldsymbol{\chi}_\mathcal{C} - \boldsymbol{\chi}_q)} A_\mathcal{C}}{\sum_{i=0}^{n} (A_i^{\text{attn}}(\boldsymbol{\chi}_\mathcal{C} - \boldsymbol{\chi}_i)^2)}, \tag{8}$$

where $\boldsymbol{\chi}_\mathcal{C}$ is the position of the control point, with $A_\mathcal{C}$ its amplitude, while $\boldsymbol{\chi}_q$ is the transducer origin. Using these resulting complex values for the transducer emissions with equation (6), we generate hypothesized single control point fields $\Psi_{\Omega\mathcal{C}}^{1,\ldots,m}$. From these fields, we construct the matrix $\mathbf{R}$ as:

$$\mathbf{R} = \begin{bmatrix} \Psi^1(\boldsymbol{\chi}_{\mathcal{C}^1}) & \ldots & \Psi^m(\boldsymbol{\chi}_{\mathcal{C}^1}) \\ \vdots & \ddots & \vdots \\ \Psi^1(\boldsymbol{\chi}_{\mathcal{C}^m}) & \ldots & \Psi^m(\boldsymbol{\chi}_{\mathcal{C}^m}) \end{bmatrix}, \tag{9}$$

such that both the matrix/eigenvector and eigenvalue/eigenvector product give a vector of the amplitudes and phases of the control points given an amplification eigenvalue. This is estimated with the assumption that the eigenvector describes a weighted 'mixing'

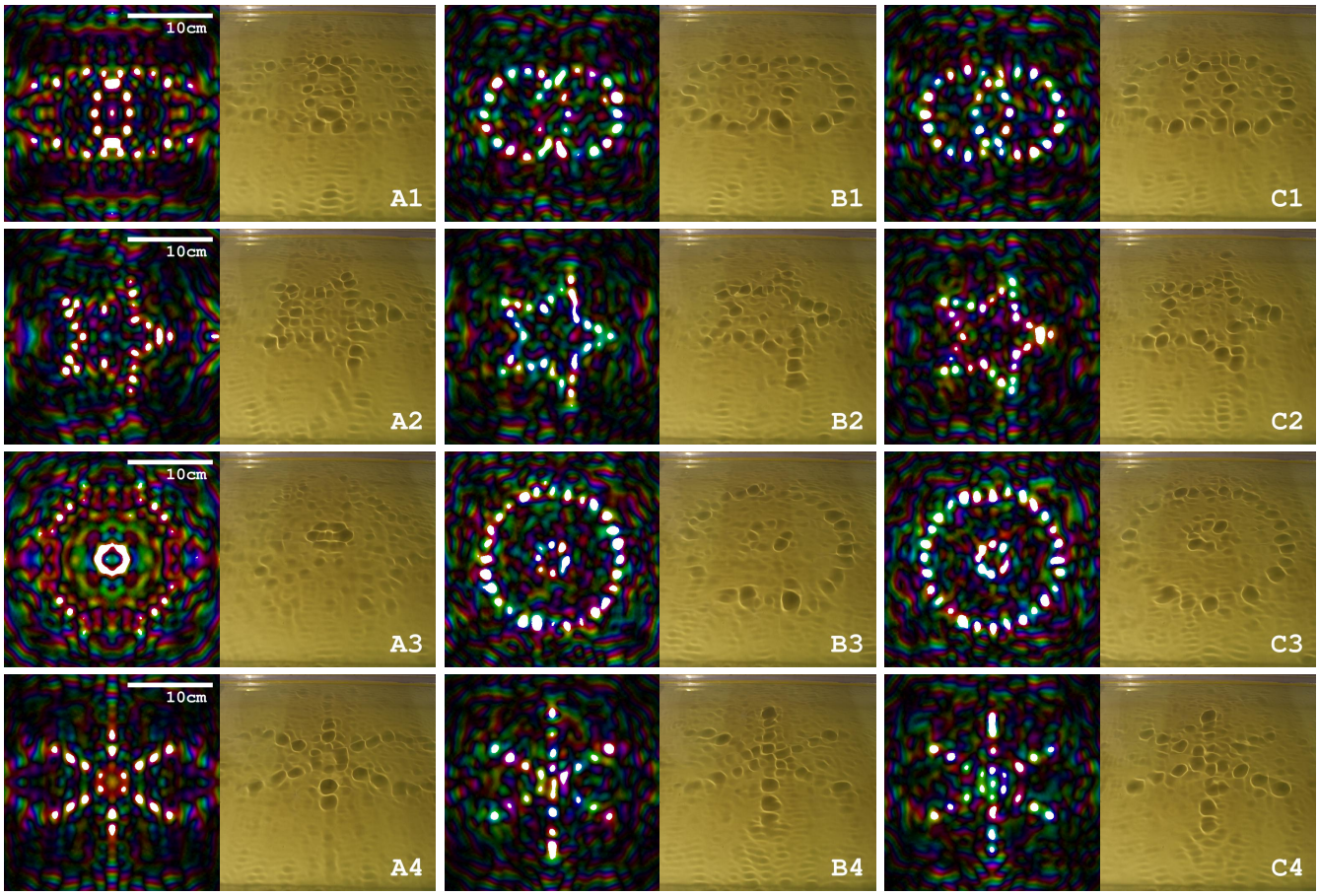

Figure 6: *Sets of shapes and boundaries as a rendered acoustic field generated by different techniques, which have then been impressed upon the surface of oil. Note that all of these figures have been generated without the use of the modulation efficiency technique as described in Section 3.4. A) Linear multiplexing approach. Some points are missing, some misplaced, with prominent secondary maxima. B) Our technique without regularization. Some points are weak, some misshapen. C) Our technique with regularization. Points are less erratic and have reduced noise. 1) Two intersecting rings, each with 16 control points, totalling 32 points overall. 2) A five-pointed star containing 30 control points. 3) Two concentric circles, wherein the outer circle is made from 24 points and the center circle is made from 8 points, with 32 points overall. 4) A six-pointed asterisk shape made up of 21 control points.*

of each single control point solution which generates maximally amplified control points (assuming the local modification effects are preserved in a global solution).

Finding a large eigenvalue λ then corresponds to a large constructive amplification of the phases in the eigenvector, which makes using these eigenvector phases desirable in any later linear system as the benefits of this amplification should generate similar amplitudes, and so more efficient transducer usage.

As only the eigenvector with the largest eigenvalue (amplification) needs to be found, a very simple power method approach can be employed. Also, as this is a step that makes the array more efficient, the method does not need to completely converge, and can be estimated and restarted from a previous solution, enabling a time-bounded solution.

This ensures that when the solution of the linear system is obtained the given control points are able to coexist while minimizing both unwanted mutual exclusion caused by destructive interference, and noise caused by constructive interference.

3.3 Weighted Tikhonov Regularization

The set of computed aperture wavefunctions for the ultrasound transducers can be seen as a linear basis set for the space of all possible interference patterns for this group of emitters, so the algorithm should find the best pattern to fit the desired output. Although we have obtained sets of control points with feasibly chosen phases, our solution method is unaware of the physical power limitations of the array. While we could follow Gavrilov and Hertzberg to optimize for the most efficient array power while not exceeding a maximum, their techniques are iterative and rely on the small numbers of control points involved producing thin matrices that are quickly decomposed. These therefore do not scale well to the large systems of control points needed for this system.

In order to find a balance of solving for the emitter configuration to a scale factor and a reasonable array efficiency, we turn to regularization techniques. Particularly the Tikhonov regularization technique is of interest because it has both the ability to constrain the solution of the linear system and so the power requirements, and has an easily specified matrix augmentation. We can augment our

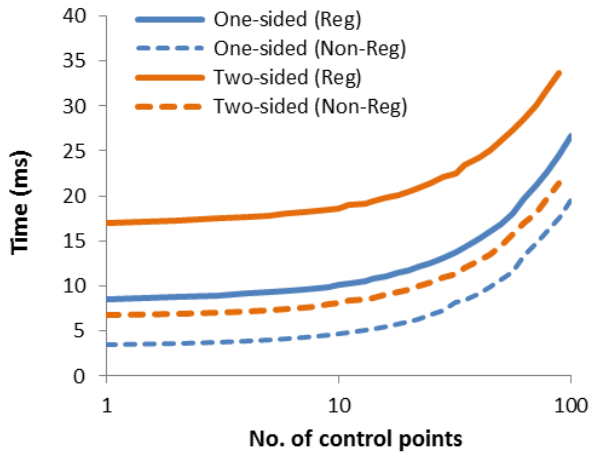

Figure 7: *The relative performance of each of our techniques, as the time taken to prepare an array update for a given number of control points randomly placed above the array. As can be seen from the graph, the system scales well at high control point counts. The regularized system is more costly to compute and the two-sided modulation takes around twice the time to compute, although this narrows when many control points are considered.*

original $\mathbf{Ax} = \mathbf{b}$ linear system from Section 3.1 as:

$$\begin{bmatrix} & \mathbf{A} & \\ \sigma_1^\gamma & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_n^\gamma \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{b} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{10}$$

where we have augmented the matrix with a block of diagonals raised to the power γ. Calculating appropriate σ_q values can be achieved with:

$$\sigma_q = \sqrt{\left\| \sum_{i=0}^{m} \frac{A_q^{\text{attn}}(\boldsymbol{\chi}_{\mathcal{C}^i} - \boldsymbol{\chi}_q) A_{\mathcal{C}^i}}{m} \right\|} \tag{11}$$

Care should be taken that $\sigma_q > 0$, so that the problem matrix remains full rank. The value γ is now chosen to be a value between zero and one. Here zero is a preference for a minimum-norm solution and can be seen as 'turning off' the regularization effect up to a scaling factor, non-integral values give fractional regularization effects. A γ-value of one then results in an attempt to counter-balance the full output power of the transducers at the control points. This solution specifies complex transducer output values that are of more equal amplitude, which results in less variance overall. Due to this, if the transducers are then powered fully, fewer artifacts are created, resulting in appreciably better output.

3.4 Modulation Efficiency

Having found a set of phases and amplitudes to use to generate an acoustic field, as previously described they can be made perceptible via a low frequency modulation. This means for example, for an 200Hz modulation the array is alternately powered for a 1/400$^{\text{th}}$ second duration and unpowered for the next 1/400$^{\text{th}}$ second. An unwanted consequence of this is a loss in power, but the cycle is necessary for the modulation to be generated.

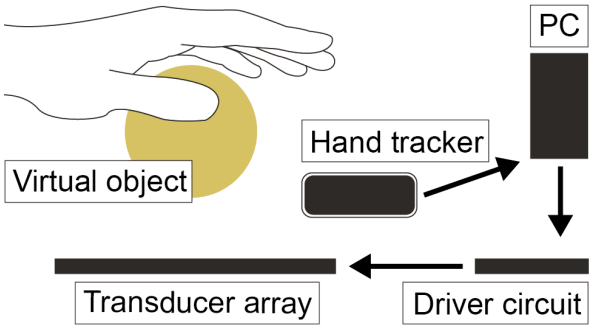

Figure 8: *The setup of our system for generating three-dimensional haptic shapes.*

We circumvent this loss by using both halves of the modulation cycle to contribute to the output. By considering the set of control points to be generated and splitting it into two sets, we can emit them in an alternating fashion, removing this shortcoming. As the phase eigenvector solution is more effective when the points are close by, we use principal component analysis to find a splitting plane to generate two control point groups that are locally dense. Then by alternating between these two acoustic field solutions, an overall perception using many control points can be generated that is almost twice as powerful as before as can be seen from Figure 4.

3.5 Algorithm Evaluation

To visually inspect the acoustic output of the technique, we used an approach in which we turn the ultrasonic array to face the surface of a thin layer of oil, as shown in Figure 5. When the ultrasound is focused, the oil surface displaces. This displacement can be enhanced by lighting the surface from a shallow angle, refracting light through the oil and revealing the structure of the acoustic field as caustics. This technique is however imperfect as the ultrasound is not completely reflected from the surface and resonance can occur causing artifacts that appear similar to noise. Lighting conditions can also cause the appearance of noise as the setup is sensitive to angle. In spite of the drawbacks, this approach is simple and effective at producing imagery from shaped acoustic fields.

To evaluate the method, we generated simulations and compared them to a visual inspection using the oil impression technique to determine the performance of our algorithm both with and without the use of regularization. For comparison we have also included the results of the linear multiplexing technique, where control points are considered singly via a time of flight calculation, summed and the complex valued transducer output normalized. This can be written as:

$$e^{i\phi_q} = \frac{\sum_{i=0}^{m} e^{2\pi i(\boldsymbol{\chi}_{\mathcal{C}^i} - \boldsymbol{\chi}_q)/\lambda}}{\left\| \sum_{i=0}^{m} e^{2\pi i(\boldsymbol{\chi}_{\mathcal{C}^i} - \boldsymbol{\chi}_q)/\lambda} \right\|} \tag{12}$$

The results of simulations using each of the three techniques are shown side-by-side for four different shapes in Figure 6.

In each case, the amplitude of each transducer is considered to be unit, with only the ultrasound phase controllable. Although this means that there is no difference in overall power, the fidelity of the results and the consistency of control points in each arrangement is markedly different.

In the algorithm comparison shown, we found that although the linear multiplexing technique was effective for points close to the

Figure 9: *Hand-object intersection sampling. a) The hand in the scene touching the virtual cube with the ultrasonic transducer array position shown below it. b) The hand is converted into sixteen planes, some of which intersect the object in the scene, cutting the hand. c) The hand-object intersections are found as line segments and processed into contiguous arcs, from which control points are derived.*

array center, it failed or gave misshapen results for points away from the center, this can be seen from Figure 6 parts A1, A3 and A4. This is due to amplification only occuring when points are close together or in the center. The differences between the normal and regularized versions of the algorithm are more subtle, for instance the missing control point in the bottom right of B1, which appears in C1, and the markedly more regular amplitude of the points in C1. The erroneous control point that appears at the lower right tip of B2, but not in C2, is a further example of the improvement that the regularization makes on top of the technique.

We then tested the update rate of the system with these control point solvers as shown in Figure 7, where we use a GeForce GTX 780 Ti graphics card to show the relative speeds of each technique. The regularisation technique was more computationally costly and the two-sided modulation solution was for small control point sets twice the solution time, but as the sets grow larger the relative differences narrow slightly.

From these results we conclude that volumetric shapes can be created using the algorithm that we have described. In the next section we present the hardware implementation and how the hand-shape intersection sampling is carried out.

4 Implementation

4.1 System Setup

We used a system consisting of an ultrasonic phased transducer array actuated by a driver circuit and a hand tracker together with a PC as shown in Figure 8. The algorithms were implemented with OpenCL on a GPU.

The ultrasound is emitted from an array of 320 piezoelectric transducers. These are driven by 5 interconnected driver boards each with two processors. The computed phase delays and amplitude values are sent from the PC to the USB controller. This consists of a USB interface and a processor. This sorts the received data and forwards it on to the processor controlling the corresponding transducers. All of the processors have synchronized clocks. They then produce one square wave output for each transducer in accordance with the phase delays and amplitude values. These output signals are amplified from 5V to 15V before leaving the driver board.

We used XMOS L1-128 processors running at 400MHz. The outputs had a refresh rate of 2MHz. We chose muRata MA40S4S transducers as they produce a large amount of sound pressure (20 Pascals at a distance of 30cm) and have a good angle of directivity (60°).

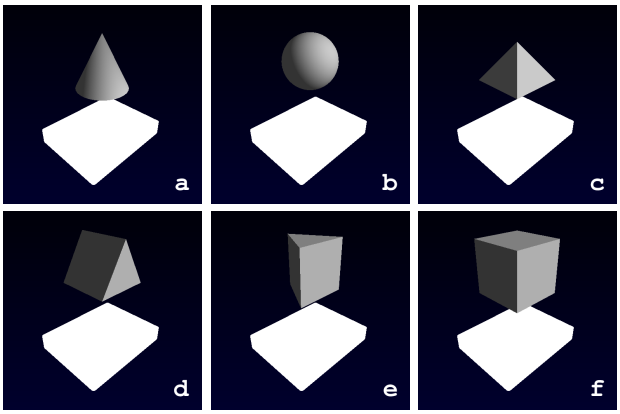

Figure 10: *The shapes involved in the user study. Shape a) was used as a training shape, before the study was carried out on shapes b) to f). Each of the shapes was scaled equally on all axes to fit a 10cm cube. The white cuboid indicates the position of the array.*

To find hand and virtual object intersections, we used a Leap Motion Controller [Leap Motion Inc. 2012]. The Leap Motion Controller has a range of 100cm and a field of view of 140 degrees, and is specialized for hand tracking making it suitable for our requirements.

4.2 Hand-Object Intersection Sampling

3D shape recognition is highly dependent on edges and vertices [Plaisier et al. 2009]. In order to create the most effective cues for shape in a volumetric space, we must generate an edge analogue to facilitate shape recognition in mid-air. To do this, we must detect the interactions between hands and shape boundaries, so that the acoustic field that we generate can create the necessary haptic feedback.

From the hand model provided by the Leap Motion Controller, we take the bone and joint positions to create the model that we use for shape boundary intersection. Each hand contains sixteen planar quadrilaterals, a palm polygon and three separate polygons for each finger.

These object-hand interactions, as shown in Figure 9, result in polygon-polygon intersections, and the resulting intersection primitive is a line segment. These are then assembled into contiguous arcs, which can be sampled to produce control points. Different sampling strategies were then employed to optionally enhance regions of high curvature.

Curvature Adaptive Parameterization

One method of modifying the sampling from a simple uniform density approach is to change the control point density dynamically by correlating control point density with curvature. This effectively increases the strength of the haptic feedback at areas of high curvature which serves to draw attention to geometrically salient features of shapes.

We determine these features by considering curvature approximations on meshes. As we are primarily dealing with small, simple meshes, we turn to the identification of local angular defects (the variation in angular sum of the surrounding triangles from 2π) in the mesh as our indicator function.

To tie this to control point density, we interpolate the curvature

Figure 11: *The user study setup.*

along mesh edges. Then at the point where the mesh and hand interact, the polygon-polygon intersection occurs, and line segment that is generated has its local curvature indicator specified by:

$$\kappa_i(t) = \kappa_0 e^{c(t_0 - t)} + \kappa_1 e^{c(t - t_1)} \tag{13}$$

where t is the parametric coordinate of the line, t_0 the point intersecting the mesh edge at the beginning of the line, t_1 the point intersecting the mesh edge at the end, c a decay constant, κ_0 the edge curvature at the beginning and κ_1 the curvature at the end. The constant c is chosen such that as the intersected straight line segment becomes longer, the curvature interpolated from the endpoints is subject to exponential decay causing the curvature sampled from the center of a flat polygon tend towards zero.

This curvature $\kappa_i(t)$ can then be used to specify local control point density, where the minimum control point density is expressed when the curvature is zero, and a higher control point density is used when the curvature is large.

5 User Study

In Section 3.5, we determined that the system is capable of producing shapes. However, as there have been no studies performed for ultrasonic haptic shapes, we need to evaluate the effectiveness of our system in conveying the shape information. Thus, we conducted an experiment where participants were asked to identify volumetric shapes by active exploration without any visual feedback in the volume above the transducer array.

We generated six volumetric shapes in total as shown in Figure 10: (a) cone, (b) sphere, (c) pyramid, (d) horizontal prism, (e) vertical prism and (f) cube. All the shapes were contained in a cubic region with 10cm sides at 17cm height centred above the transducer array. The region containing shapes therefore extends between 12cm and 22cm above the array surface. We decided to use 15mm arc length per control point as this gives a good trade-off between actuated volume and number of control points for a uniform distribution.

From our trial runs, we found that it was easy to confuse the cone and the pyramid due to their similar footprints during hand-shape intersections. Without any visual cues, it is difficult to identify the exact location of parts of the shape resulting in misinterpretation of the haptic feedback. Thus, the cone was used as a practice shape. Each participant was shown a printout of all the six shapes as a list of choices. Participants were instructed to explore the area of above the array and asked to guess which shape that they thought it was among the six shapes. As this was the practice, they were informed whether or not they were correct. After that, they were asked to explore the cone for a few more minutes until they were confident that they could identify the shape well.

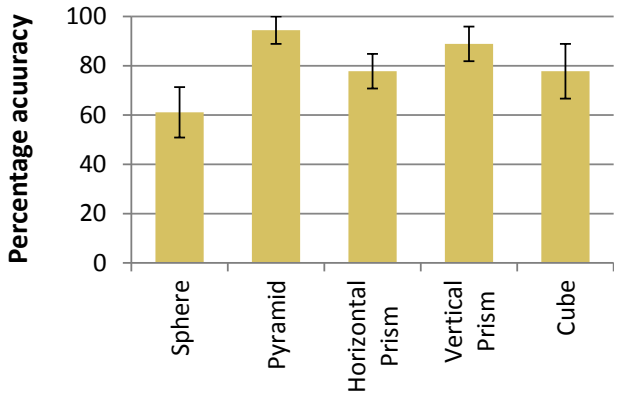

Figure 12: *Percentage accuracy across all the participants for each shape. Error bars denote +/- standard error of the mean*

Once the practice was completed, the actual tests were run with the five other shapes i.e. (b) to (f). Each of the five shapes was repeated three times for a total of 15 trials and the trials were presented in a randomized order. Participants were informed that the cone will not be included in the test and were not told if their answers were correct after each trial. Additionally, the participants were not informed on the number of trials nor if all the shapes will be present. They also wore headphones that played white noise to mask any audible cues from the system. Figure 11 shows the setup of the user study.

Results

Six participants (all males with aged range 27 to 35) took part in the user study which lasted about 25 minutes. They were all coincidentally right-handed. None of them have previously performed any studies involving ultrasonic haptic shapes.

The performance of each participant in correctly identifying all the shapes ranged from 66.7% to 100% (mean 80, SD 12.6). Figure 12 shows the percentage accuracy for each shape across all the participants (out of 18 trials). The pyramid was the easiest to identify (94.4%, SEM 5.6%) and the sphere the least (61.1%, SEM 10.2%). Table 1 showed which shapes were more likely to be confused by the participants. The pyramid was never confused with either the horizontal prism, vertical prism or the cube. The horizontal prism was never confused with the vertical prism or the sphere. The vertical prism was never confused with the cube or the pyramid.

	Sphere	Pyramid	Horizontal Prism	Vertical Prism	Cube
Sphere	11	2		4	1
Pyramid	1	17			
Horizontal Prism		3	14		1
Vertical Prism	1		1	16	
Cube	1		2	1	14

Table 1: *Confusion matrix showing the shapes that were most frequently confused across all the participants (out of 18 trials for each shape)*

There are perceptual issues associated with the localisation of ultrasonic haptic feedback with the lack of visual feedback which would have resulted in the misidentification of the shapes [Wilson et al. 2014; Hoshi et al. 2010]. Research has shown that exploratory strategies such as hand motion, contour following or one-finger/one-hand can also affect information derived from an object's shape [Lederman and Klatzky 1987]. In this study, the participants were never informed of any technique to help them identify a shape.

As can be seen from the results, even with very little training and using naive users, the participants were generally successful at identifying the shapes without any visual feedback. Overall, the results demonstrate that our system is efficient in rendering perceivable haptic shapes.

6 Applications

Inaccessible objects such as those in museum cases or inside the human body, can be visually explored through bi-directional mirrors or neurosurgical props. While these methods allow the user to interact with the objects and intersect them with their hands, they offer no haptic feedback. Augmenting with our system enables superior spatial learning and the ability to highlight valuable information through haptic feedback. Figure 13 (left) depicts a surgeon exploring a CT scan with haptic feedback allowing them to feel tumors.

Touchless Interfaces are becoming increasingly common. They afford an intuitive, flexible user interface but lack the haptic feedback provided by physical buttons and controls. In many situations, such as in the cockpit of a vehicle (see Figure 13 (center)), the user will want to operate the system while their eyes are busy on another task. Integrating our system would allow the user to feel the geometry of an interface and localize on a specific item.

Virtual Reality has long been a goal of interactive systems. Haptic feedback provides our sense of proprioception, kinesthesia and touch making it essential for an effective system. Recent advances in head mounted displays have greatly improved the realism of the visual feedback, yet haptic feedback still requires proxy or wearable haptic devices. This unnatural disconnection breaks the immersion of a virtual reality. Our system would enable users to freely explore the virtual world unencumbered while receiving haptic feedback from the objects that they interact with, as shown in Figure 13 (right).

7 Limitations and Future Work

We have demonstrated a system for producing volumetric haptic shapes in mid-air. We have also shown that users are capable of discriminating and identifying volumetric haptic shapes with high accuracy. Nonetheless, the ultrasonic haptic feedback technology has some drawbacks.

Firstly, the shape created must be in the working volume of the device for it to function correctly. Too far from the device, or moving the shapes out of the working volume, and the device loses power, being able to focus less and less as the users' hands moves outside of the active region. As the pressure drops off as the reciprocal of distance from the transducer, this implies that the intensity and power obey an inverse square law with distance. We find that currently our system peforms best between 15cm and 50cm.

Secondly, our system only functions with hands. Other parts of the human body can have more difficulty detecting a particular frequency of vibrations generated by the device or are less sensitive to vibration generally. Given this, our system implementation is limited to hand-based haptics.

Thirdly, we are limited by the size and power of the transducer array in the number and strength of control points. Increasing the number of control points reduces the strength of all control points with a fixed number of transducers. This implies that there is an important trade-off between the point sampling density and rendering quality when generating shapes. While this is in part ameliorated by only rendering the hand-object intersections to decrease the control point count, the problem remains. In this light, the balance between both properties in which the shape identification is most efficient is yet to be investigated.

For future work, we plan to investigate this issue to determine where the optimal balance between strength and rendering quality lies. By exploiting this point, further increases in shape complexity and subtlety could be realized.

8 Conclusions

This paper demonstrated that our algorithm can control the volumetric distribution of the acoustic radiation force field when using a two-dimensional phased ultrasound array. By controlling the acoustic field, we are able to produce a volumetric haptic shapes in mid-air. This algorithm is capable of running interactively, allowing us to create real-time haptic sensations and dynamically changing shapes if we wanted to.

The capability to create mid-air haptic shapes is advantageous as it allows the user to 'walk-up and use' the system. The user does not need to wear any attachments or be restricted by the use of any tools thus encouraging spontaneous use and allowing freedom to explore. This is the first algorithm that we know of that creates and controls feelable three-dimensional shapes using a volumetric distribution of ultrasound.

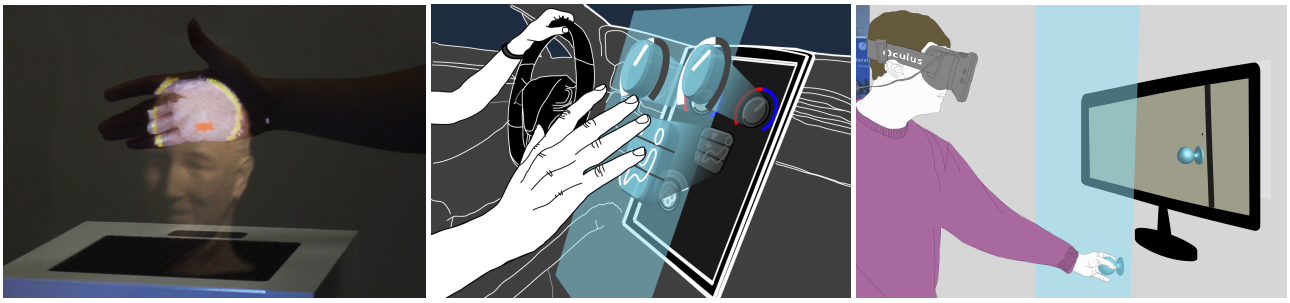

Figure 13: *Left: Our system adds haptic information to the results of a CT scan. Center: Augmenting the dashboard of a car with our system enables eye-busy interactions. Right: Opening a door to a room filled with monsters in a game becomes more immersive with our system.*

Acknowledgements

We thank Matt Sutton for taking the images and the video. This work is supported by European Research Council (Starting Grant Agreement 278576) under the Seventh Framework Programme.

References

ALEXANDER, J., MARSHALL, M. T., AND SUBRAMANIAN, S. 2011. Adding haptic feedback to mobile tv. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, ACM, New York, NY, USA, CHI EA '11, 1975–1980.

BERKELMAN, P. J., BUTLER, Z. J., AND HOLLIS, R. L. 1996. Design of a hemispherical magnetic levitation haptic interface device. In *in Proceedings of the ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 17–22.

BIANCHI, G., KNOERLEIN, B., SZÉKELY, G., AND HARDERS, M. 2006. High precision augmented reality haptics. In *Proc. EuroHaptics*, 169–178.

CARTER, T., SEAH, S. A., LONG, B., DRINKWATER, B., AND SUBRAMANIAN, S. 2013. Ultrahaptics: multi-point mid-air haptic feedback for touch surfaces. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, ACM, 505–514.

DALECKI, D., CHILD, S., RAEMAN, C., AND CARSTENSEN, E. 1995. Tactile perception of ultrasound. *The Journal of the Acoustical Society of America 97*, 3165.

DIPIETRO, L., SABATINI, A. M., AND DARIO, P. 2008. A survey of glove-based systems and their applications. *Trans. Sys. Man Cyber Part C 38*, 4 (July), 461–482.

EBBINI, E. S., AND CAIN, C. A. 1989. Multiple-focus ultrasound phased-array pattern synthesis: optimal driving-signal distributions for hyperthermia. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 36*, 5, 540–548.

GAVRILOV, L., AND TSIRULNIKOV, E. 2002. Mechanisms of stimulation effects of focused ultrasound on neural structures: Role of nonlinear effects. *Nonlinear Acoustics at the Beginning of the 21st Century 1*, 445–448.

GAVRILOV, L., AND TSIRULNIKOV, E. 2012. Focused ultrasound as a tool to input sensory information to humans (review). *Acoustical Physics 58*, 1–21.

GAVRILOV, L. 2008. The possibility of generating focal regions of complex configurations in application to the problems of stimulation of human receptor structures by focused ultrasound. *Acoustical Physics 54*, 269–278.

GESCHEIDER, G., AND WRIGHT, J. 2008. *Information-processing channels in the tactile sensory system: A psychophysical and physiological analysis*. Psychology Press.

HAYWARD, V., ASTLEY, O., CRUZ-HERNANDEZ, M., GRANT, D., AND ROBLES-DE-LA-TORRE, G. 2004. Haptic interfaces and devices. *Sensor Review 24*, 1, 16–29.

HERTZBERG, Y., NAOR, O., VOLOVICK, A., AND SHOHAM, S. 2010. Towards multifocal ultrasonic neural stimulation: pattern generation algorithms. *Journal of neural engineering 7*, 5, 056002.

HOSHI, T., TAKAHASHI, M., IWAMOTO, T., AND SHINODA, H. 2010. Noncontact tactile display based on radiation pressure of airborne ultrasound. *IEEE Transactions on Haptics 3*, 3, 155 –165.

LEAP MOTION INC, 2012. Leap Motion Controller. `http://leapmotion.com`.

LEDERMAN, S. J., AND KLATZKY, R. L. 1987. Hand movements: A window into haptic object recognition. *Cognitive psychology 19*, 3, 342–368.

NASCOV, V., AND LOGOFĂTU, P. C. 2009. Fast computation algorithm for the Rayleigh-Sommerfeld diffraction formula using a type of scaled convolution. *Appl. Opt. 48*, 22 (Aug), 4310–4319.

PLAISIER, M. A., BERGMANN TIEST, W. M., AND KAPPERS, A. M. L. 2009. Salient features in 3-d haptic shape perception. *Attention, Perception, & Psychophysics 71*, 2, 421–430.

REINER, M. 2004. The role of haptics in immersive telecommunication environments. *IEEE Transactions on Circuits and Systems for Video Technology 14*, 3 (march), 392–401.

SODHI, R., POUPYREV, I., GLISSON, M., AND ISRAR, A. 2013. Aireal: interactive tactile experiences in free air. *ACM Transactions on Graphics (TOG) 32*, 4, 134.

SUZUKI, Y., AND KOBAYASHI, M. 2005. Air jet driven force feedback in virtual reality. *Computer Graphics and Applications, IEEE 25*, 1 (Jan), 44–47.

WILSON, G., CARTER, T., SUBRAMANIAN, S., AND BREWSTER, S. A. 2014. Perception of ultrasonic haptic feedback on the hand: Localisation and apparent motion. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, ACM, New York, NY, USA, CHI '14, 1133–1142.

WUSHENG, C., TIANMIAO, W., AND LEI, H. 2003. Design of data glove and arm type haptic interface. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2003. HAPTICS 2003. Proceedings. 11th Symposium on*, 422 – 427.